Software Application Profile: mrrobust - a tool for performing two-sample summary Mendelian randomization analyses

Wes Spiller[1], Neil M. Davies[1]and Tom M. Palmer[2]

[1] Population Health Science Institute, University of Bristol, Barley House, Oakfield Grove, Bristol, BS8 2BN, U.K [2] Department of Mathematics and Statistics, Lancaster University, Lancaster, United Kingdom

Abstract

Motivation: In recent years Mendelian randomization analysis using summary data from genome-wide association studies has become a popular approach for investigating causal relationships in epidemiology. The mrrobust Stata package implements several of the recently developed methods.

Implementation: mrrobust is freely available as a Stata package.

General Features: The package includes inverse variance weighted estimation, as well as a range of median, modal and MR-Egger estimation methods. Using mrrobust, plots can be constructed visualising each estimate either individually or simultaneously. The package also provides statistics such as I_{GX}^2, which are useful in assessing attenuation bias in causal estimates.

Availability: The software is freely available from GitHub [https://raw.github.com/remlapmot/mrrobust/master/].

Key words: summary MR, MR-Egger, IVW, weighted median, Mendelian randomization, Stata

---

Key Messages

- The mrrobust software package facilitates two-sample summary MR analyses using summary data from genome-wide association studies.

- The package allows for implementation of a range of summary MR estimators using Stata, improving the extent to which results are reproducible.

- Conclusions from the supported analyses can be robust to sources of confounding bias and pleiotropy, though findings should be considered with respect to the underlying assumptions of each estimator.

Introduction

Mendelian randomization(1) has developed into a popular approach to examining causal relationships in epidemiology(2, 3). By employing genetic variants as instrumental variables (IVs) it is possible to limit bias from confounding, provided variants satisfy the assumptions of IV analysis(1, 4). For a genetic variant to serve as a suitable instrument, three assumptions must hold, 1) it must be associated with the exposure of interest, 2) there must be no confounders of the instrument and outcome, and 3) the instrument must not affect the outcome except via the exposure of interest.(5).

Candidate variants are usually identified through large genome-wide association studies (GWASs) (6). However, IV analyses using single variants rarely have sufficient power to test hypotheses of interest(6, 7). One approach to increase the statistical power of Mendelian randomization studies is to use multiple genetic variants as instruments within a two-sample summary framework(8, 9). Two-sample Mendelian randomisation estimates the effect of the exposure using instrument-exposure and instrument-outcome associations from different samples, often through methods originally developed for meta-analysis(8, 9). This is particularly useful, as MR estimators such as MR Egger and median based regression are robust to certain forms of violation of the third instrumental variable assumption(8, 10, 11). Violations of this assumption can occur through directional pleiotropy- where a genetic variant affects the study outcome through pathways that are not mediated via the exposure. Such developments have contributed to the increasing popularity of two-sample summary MR(5).

This paper introduces the mrrobust Stata package as a tool for performing two-sample summary MR analyses. The mrrobust package is a tool to help researchers implement two-sample MR analyses, and can be viewed as the Stata counterpart to toolkits such as the MR-Base web application, and the MendelianRandomization and TwoSampleMR R packages(12, 13). Whilst it is possible to conduct individual level IV analyses in Stata using modules such as IVREG2(14), two-sample summary MR has previously required bespoke code to implement. The mrrobust package addresses this limitation, providing a suite of popular two-sample MR methods and sensitivity analyses. Before continuing, we briefly outline the three primary estimation methods included in the mrrobust package, using the notation of Bowden et al(10, 15).

Inverse variance weighting (IVW)

To perform IVW a weighted average $\hat{\beta}_{IVW}$ is calculated using the set of ratio estimates $\hat{\beta}_J$ for each individual variant $J = 1,2,...,j$ (9). Ratio estimates are obtained for each variant by dividing the instrument-outcome association by the corresponding instrument-exposure association. Such association estimates are obtained by fitting simple linear regression models of the outcome and exposure upon the genetic variant, primarily by conducting a GWAS. Let $\hat{\gamma}_j$ and σ_{Yj}^2 denote the instrument-outcome association and variance respectively for the jth variant. The IVW estimate is then defined as:

$$\hat{\beta}_{IVW} = \frac{\sum_{j=1}^{J} w_j \hat{\beta}_j}{\sum_{j=1}^{J} w_j}, w_j = \frac{\hat{\gamma}_j^2}{\sigma_{Yj}^2}$$

This corresponds to the estimate one would obtain from a weighted linear regression of the set of instrument-outcome associations upon the set of instrument-exposure associations, constraining the intercept at the origin(9). One drawback of the IVW approach is that causal effect estimates can be biased in cases where one or more variants exhibit directional pleiotropy(9).

MR-Egger regression

MR-Egger regression is valid under weaker assumptions than IVW, as it can provide unbiased causal effect estimates even if the variants have pleiotropic effects.

In this case, the set of instrument-outcome associations is regressed upon the set of instrument-exposure associations, weighting the regression using precision of the instrument-outcome associations as in the IVW case(8). However, MR-Egger does not constrain the intercept at the origin, and the intercept represents an estimate of the average directional pleiotropic effect across the set of variants. The slope of the model provides an unbiased estimate of the causal effect(8, 10). If there is little evidence of systematic differences between the IVW and MR-Egger, then the IVW should be preferred. The IVW is more efficient, but potentially less robust, and in such cases the IVW estimate is often most appropriate estimate to adopt due to the greater precision of IVW estimates in comparison with other approaches(10). If there are differences between the IVW and MR-Egger estimates, this may be due to pleiotropy or heterogeneous treatment effects.

The utility of MR Egger regression hinges upon two core assumptions. First, the INstrument Strength Independent of Direct Effect (InSIDE) assumption requires the effects of SNPs on the exposure and their pleiotropic effects on the outcome to be independent. If the InSIDE assumption holds, estimates for variants with stronger instrument-exposure associations ($\hat{\gamma}_j$) will be closer to the true causal effect parameter than variants with weaker associations(8). Second, the NO Measurement Error (NOME) assumption requires no measurement error to be present in the instrument-exposure associations, and therefore that the variance of the instrument-exposure association $\sigma_{Xj}^2 = 0$. In cases where NOME is strictly satisfied, estimates $\hat{\gamma}_j$ will be equal to γ_j and the variance of the ratio estimate for each variant j is $var(\hat{\beta}_j) = \frac{\sigma_{Yj}^2}{\hat{\gamma}_j^2}$. We further note that the NOME assumption applies to other two sample MR approaches and is not therefore a unique feature of the MR Egger approach.

In cases where the NOME assumption is violated, individual variants will suffer from weak instrument bias, leading to attenuation of MR Egger estimates towards the null. This can occur if the SNPs were not genome-wide significant ($p = 5 \times 10^{-8}$) or were selected from small GWAS. One novel approach to assessing the strength of the NOME assumption is to evaluate the I_{GX}^2 statistic, interpreted as the relative degree of attenuation bias in the MR Egger regression in the interval (0,1)(10). Thus, for example, an I_{GX}^2 value of 0.7 represents an estimated relative bias of 30% towards the null. Further details regarding calculation of the I_{GX}^2 statistic are presented in the supplementary material.

Weighted median

The weighted median approach is an adaptation of the simple median estimator for two-sample summary MR(15). For a total number of variants $J = 2k + 1$, the simple median approach selects the middle ratio estimate $\hat{\beta}_{k+1}$, from ordered ratio estimates $\hat{\beta}_1, \hat{\beta}_2, \ldots \hat{\beta}_j$(15). In cases where the total number of variants is even, the median is interpolated as $\frac{1}{2}(\hat{\beta}_k + \hat{\beta}_{k+1})$. As the simple median approach is inefficient, particularly in cases with variable precision in the set of ratio estimates, it is preferable to incorporate weights in a similar fashion to the IVW and MR Egger approaches. Let $s_j = \sum_{k=1}^{j} w_k$ be the sum of weights for the set of variants $1,2,\ldots j$, standardised so the sum of weights s_J=1. The weighted median estimator is the median of a distribution having estimate $\hat{\beta}_j$ as its $p_j =$

$100 \left(s_j - \frac{w_j}{2} \right)^{th}$ percentile(15). For the range of percentile values, we perform a linear extrapolation between neighbouring ratio estimates.

An important assumption of the median summary MR approaches is that more than 50% of the genetic variants do not exhibit directional pleiotropy. In the simple median case, this threshold refers to the number of variants, whilst in the weighted median case the 50% threshold is with respect to the weights of the non-pleiotropic variants(15).

Additional estimators

As two-sample MR represents a developing area of genetic epidemiology, novel approaches to causal effect estimation are incorporated into the mrrobust package through frequent updates. One such method is the mode-based estimator put forward by Hartwig et al(16). Details on the implementation of this approach with accompanying examples can be found in the supplementary material.

Visualizing MR estimates

One useful approach to presenting the results of MR analyses is to produce a scatterplot, with the x and y axes representing the instrument-exposure and instrument-outcome associations respectively for each variant. If one were to draw a hypothetical regression line leading from the origin to each variant, the slope of the line would represent a ratio estimate of the causal effect using the single variant as an instrument, that is, dividing the instrument-outcome association by the instrument-exposure association (defined as β_j above). The precision of the instrument-outcome association estimate for each variant is illustrated using vertical error bars, whilst horizontal error bars pertaining to the instrument-exposure association may be omitted for clarity. As the IVW, MR-Egger, median, and modal approaches essentially meta-analyse the set of ratio estimates, it is possible to include regression lines highlighting effect estimates of each approach for comparison. For such regression lines, positive and negative slopes are indicative of a positive or negative effect respectively, whilst a slope of zero represents the absence of an observed association.

Implementation

The mrrobust package uses functions from moremata(17), addplot(18), and the heterogi(19) command. For versions of Stata 13 and higher, it can be installed using the .net install

command from [https://raw.github.com/remlapmot/mrrobust/master/]. For older versions of Stata, a zip archive of the files is freely available for download at: [https://github.com/remlapmot/mrrobust].

The package facilitates two-sample summary MR analyses with key features including:

- IVW and MR-Egger regression approaches, including fixed effects MR-Egger regression, standard error correction, and weighting options.
- Unweighted, weighted and penalized weighted median IV estimators, providing pleiotropy robust estimates in cases where fewer than 50% of the genetic instruments are valid.
- Modal estimation following Hartwig et al(16), including weighted and unweighted variations.
- Presentation of heterogeneity statistics, statistics such as I_{GX}^2 for use in assessing attenuation bias(10), and Simulation Extrapolation (SIMEX) correction following Bowden et al(20).
- Plotting tools to visualise IVW, MR-Egger and weighted median estimators, as well as density plotting with respect to implementing the modal estimator.
- Illustrative examples and documentation using data from Do et al(21).

Applied Examples: Adiposity and Height as predictors of serum glucose levels

To illustrate key features of the mrrobust package, we perform two analyses investigating potential relationships between adiposity, height, and serum glucose. Adiposity was selected owing to the vast body of evidence supporting a positive association with serum glucose levels(22-25), whilst height was based upon limited evidence of association(26-28). Glucose was selected as an outcome with respect to its hypothesised role in the development of Type-II diabetes(22, 28). Datasets were obtained from the MR-Base web application and pruned for linkage disequilibrium prior to conducting the analyses(13).

Applied Example I: Adiposity and Serum Glucose

Though the relationship between adiposity and glucose has received much attention in the literature, such studies are predominantly observational and therefore may be subject to bias from confounding. This provides motivation for considering Mendelian randomization techniques which are able to control for such unobserved confounding. In the initial analysis,

6

we select adiposity as an exposure measured using standardised body mass index (BMI), obtaining estimates of its associations with genotypes and their respective standard errors from Locke et al(29).

For the outcome, we consider log transformed measures of serum glucose $log(mM)$ utilising effect estimates and standard errors from Shin et al(30). The summary data utilised for this analysis is provided in the supplementary material. Adopting a GWAS significance p-value threshold of 5×10^{-8} a total of 79 independent SNPs were identified in both samples. We confirmed the linkage equilibrium (LD) between the SNPs using a clumping algorithm, and a clumping distance of 10000kb, and an LD R^2 of 0.001. This resulted in a total of 79 SNPs for use as instrumental variables, details of which are presented in the supplementary material.

Using mrrobust, we conducted IVW, MR-Egger, and weighted median regression approaches using the above summary data. The code for our analysis is in the supplementary material. For IVW and MR Egger approaches the regression was weighted using the variance of the instrument-outcome association. The set of summary MR estimates are presented in Table 1A.

We find strong evidence of a positive association between BMI and serum glucose using both IVW and weighted median methods. Considering the MR Egger case, a substantial average directional pleiotropic effect was not detected, and the lack of significance with respect to the effect estimate can be attributed to a lack of statistical power. An I_{GX}^2 value of 0.88 was reported, which can be interpreted as a relative bias in the MR-Egger estimate of 12% towards the null. The estimates are shown in Figure 1A, constructed using the mreggerplot command, which generates a scatterplot of the instrument-exposure and instrument-outcome associations for each variant. This shows the set of estimates to be in agreement, with the plot being constructed as previously described.

Applied Example II: Height and Serum Glucose

As a further example, we consider the effect of standardised height (*meters*) upon serum glucose using summary data from Wood et al(31), and outcome summary data on log transformed serum glucose from Shin et al(30). The summary data utilised for this analysis is provided in the supplementary material. We assess the SNPS for LD using criteria from the

previous example and identify 367 SNPs as suitable instruments for the analysis, details of which are presented in the supplementary material. The set of summary MR estimates are presented in Table 1B.

From Table 1B we find no evidence against the null hypothesis of no association between height and serum glucose levels using IVW, weighted median, and MR Egger regression. Considering the MR Egger case, there appeared to be no evidence of directional pleiotropy, with an I_{GX}^2 value of 0.90 indicating a relative bias of 10% towards the null. As in the previous example, a plot of the MR estimates can be generated using the mreggerplot command as shown in Figure 1B. In this scenario the estimates appear in agreement, indicating a lack of evidence for a substantial directional pleiotropic effect.

Discussion

The mrrobust package is a freely available Stata package, containing a number of summary MR estimation methods which can be used to estimate causal effects. In the applied example, the mrrobust package was able to provide a series of estimates, finding evidence of a positive association between BMI and serum glucose, and no evidence of association between height and serum glucose. One possible conclusion that can be drawn from these results is that previously reported associations between height and glucose are driven by confounding factors(32, 33). It is important, however, to consider the extent to which Mendelian randomization is appropriate for a given analysis, and by extension situations in which mrrobust is suitable.

In the first instance, Mendelian randomization studies only produce unbiased estimates when genetic instruments satisfy the assumptions of each estimator (e.g. IVW, MR-Egger, or weighted median). In two-sample analyses genetic instruments should be associated with the exposure of interest at genome-wide levels of significance (satisfying the first instrumental variable assumption), and pruned for LD to limit the overlap between SNPs. The IVW estimator also requires that genetic variants should not have directional pleiotropic effects. The MR Egger and median estimators are robust to directional pleiotropy if the effects of the exposure are constant. MR Egger regression requires the InSIDE assumption, whilst median methods assume that the number of valid instruments being greater than 50%. For MR-Egger estimation where the value of I_{GX}^2 is low, it is possible to use SIMEX to correct for regression attenuation towards the null. This is implemented using the mreggersimex command.

In this paper, we have presented the mrrobust Stata package as an accessible toolkit for performing summary MR and instrumental variable analysis using many instruments. It contains a range of summary MR approaches, and should make examining causal relationships using Mendelian randomization more accessible for genetic epidemiologists.

Supplementary Data

Within the supplementary material we include example code and Stata output for each of the analyses performed within this paper, as well as the summary data obtained from the MRBase GWAS catalogue. We also include a brief summary of the I_{GX}^2 statistic, as well as guidance on implementing and interpreting the modal estimator.

Further details and supplementary materials can be found at:

References

1. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32(1):1-22.
2. Burgess S, Timpson NJ, Ebrahim S, Davey Smith G. Mendelian randomization: where are we now and where are we going? Int J Epidemiol. 2015;44(2):379-88.
3. Evans DM, Davey Smith G. Mendelian Randomization: New Applications in the Coming Age of Hypothesis-Free Causality. Annu Rev Genomics Hum Genet. 2015;16:327-50.
4. Scholder SVK, Davey Smith G, Lawlor DA, Propper C, Windmeijer F. Mendelian Randomization: The Use of Genes in Instrumental Variable Analyses. Health Econ. 2011;20(8):893-6.
5. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res. 2015.
6. Davies NM, von Hinke Kessler Scholder S, Farbmacher H, Burgess S, Windmeijer F, Davey Smith G. The many weak instruments problem and Mendelian randomization. Stat Med. 2015;34(3):454-68.
7. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. Int J Epidemiol. 2013;42(4):1134-44.

8.   Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44(2):512-25.
9.   Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37(7):658-65.
10.   Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. Int J Epidemiol. 2016.
11.   Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med. 2017.
12.   Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol. 2017.
13.   Hemani G, Zheng J, Wade KH, Laurin C, Elsworth B, Burgess S, et al. MR-Base: a platform for systematic causal inference across the phenome using billions of genetic associations. bioRxiv. 2016.
14.   Baum C, Schaffer M, Stillman S. IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation. 2016.
15.   Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016;40(4):304-14.
16.   Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol. 2017;46(6):1985-98.
17.   Jann B. MOREMATA: Stata module (Mata) to provide various functions. Statistical Software Components. 2005.
18.   Jann B. A note on adding objects to an existing twoway graph. Stata Journal. 2015;15(3):751-5.
19.   Orsini N HJ, Bottai M, Buchan I. . heterogi: Stata module to quantify heterogeneity in a meta-analysis. . Statistical Software Components. 2006.
20.   Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. Int J Epidemiol. 2016;45(6):1961-74.
21.   Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. Nat Genet. 2013;45(11):1345-+.
22.   Boden G. Role of fatty acids in the pathogenesis of insulin resistance and NIDDM. Diabetes. 1997;46(1):3-10.
23.   Kim SH, Reaven G. Obesity and insulin resistance: an ongoing saga. Diabetes. 2010;59(9):2105-6.
24.   O'Gorman DJ, Karlsson HK, McQuaid S, Yousif O, Rahman Y, Gasparro D, et al. Exercise training increases insulin-stimulated glucose disposal and GLUT4 (SLC2A4) protein content in patients with type 2 diabetes. Diabetologia. 2006;49(12):2983-92.
25.   Hjerkind KV, Stenehjem JS, Nilsen TI. Adiposity, physical activity and risk of diabetes mellitus: prospective data from the population-based HUNT study, Norway. BMJ Open. 2017;7(1):e013142.
26.   Ogonowski J, Miazgowski T. Are short women at risk for gestational diabetes mellitus? Eur J Endocrinol. 2010;162(3):491-7.
27.   Brown DC, Byrne CD, Clark PM, Cox BD, Day NE, Hales CN, et al. Height and glucose tolerance in adult subjects. Diabetologia. 1991;34(7):531-3.
28.   Asao K, Kao WH, Baptiste-Roberts K, Bandeen-Roche K, Erlinger TP, Brancati FL. Short stature and the risk of adiposity, insulin resistance, and type 2 diabetes in middle age: the Third National Health and Nutrition Examination Survey (NHANES III), 1988-1994. Diabetes Care. 2006;29(7):1632-7.

29.   Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Felix R, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518(7538):197-U401.
30.   Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. Nat Genet. 2014;46(6):543-50.
31.   Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet. 2014;46(11):1173-86.
32.   Olatunbosun ST, Bella AF. Relationship between height, glucose intolerance, and hypertension in an urban African black adult population: a case for the "thrifty phenotype" hypothesis? J Natl Med Assoc. 2000;92(6):265-8.
33.   Rehunen SKJ, Kautiainen H, Eriksson JG, Korhonen PE. Adult height and glucose tolerance: a re-appraisal of the importance of body mass index. Diabet Med. 2017.

Tables

Table 1: Summary MR estimates for the effect of standardised BMI (A) and height (B) upon log transformed serum glucose.

	Estimate	*SE*	*p-value*	*95% CI*
BMI (A)				
IVW				
Effect	0.023	*0.008*	0.004	0.01, 0.04
MR Egger				
Intercept	0.000	*0.001*	0.948	-0.001,0.001
Effect	0.022	0.022	0.325	-0.02, 0.07
Weighted Median				
Effect	0.034	*0.012*	0.005	0.01, 0.06
Height (B)				
IVW				
Effect	0.002	*0.003*	0.641	-0.005, 0.008
MR Egger				
Intercept	0.0001	*0.0003*	0.627	-0.0001,0.0001
Effect	0.003	0.009	0.777	-0.02, 0.02
Weighted Median				
Effect	<0.0001	*0.005*	>0.99	-0.01, 0.01